Identification of Linear State-Space Models Subject to Truncated Gaussian Disturbances

Identification of Linear State-Space Models Subject to Truncated Gaussian Disturbances

R. A. González, A. L. Cedeño, K. Tiels and T. Oomen

Identification under bounded disturbances

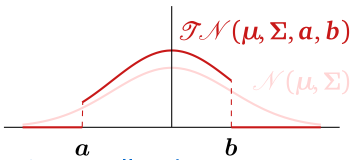
- Constraints are integral to state estimation for process optimization, monitoring, and control
- Yet, Gaussian assumptions are prevalent in identification ⇒ inaccurate for bounded-support noise

Problem formulation

Discrete-time SISO system in state-space form:

$$\begin{aligned} \boldsymbol{x}_{t+1} &= \boldsymbol{A}\boldsymbol{x}_t + \boldsymbol{B}u_t + \boldsymbol{w}_t \\ y_t &= \boldsymbol{C}\boldsymbol{x}_t + Du_t + v_t \end{aligned} \iff \boldsymbol{\xi}_t = \boldsymbol{\Theta}\boldsymbol{\zeta}_t + \boldsymbol{\eta}_t$$

- Truncated Gaussian noise $\boldsymbol{\eta}_t \sim \mathcal{TN}(\boldsymbol{\mu},\boldsymbol{\Sigma},\boldsymbol{a},\boldsymbol{b})$
- Initial condition $\boldsymbol{x}_1$ known or $\boldsymbol{x}_1 \sim \mathcal{TN}(\boldsymbol{e},\boldsymbol{H},\boldsymbol{a}_1,\boldsymbol{b}_1)$
- Unknowns: system matrices $\boldsymbol{\Theta}$, noise parameters $\boldsymbol{\mu},\boldsymbol{\Sigma},\boldsymbol{a},\boldsymbol{b}$, and possibly $\boldsymbol{e},\boldsymbol{H},\boldsymbol{a}_1,\boldsymbol{b}_1$ as well



How can we estimate all unknown parameters (stacked in $\boldsymbol{\beta}$) using $\{u_t, y_t\}_{t=1}^N$?

EM for truncated Gaussian identification

Direct maximum likelihood estimation ($\arg\max_{\boldsymbol{\beta}} \mathcal{L}(\boldsymbol{\beta})$) is intractable. Instead, we consider the EM algorithm:

1. For fixed $\hat{\boldsymbol{\beta}}^{(k)}$, obtain $\mathcal{Q}(\boldsymbol{\beta},\hat{\boldsymbol{\beta}}^{(k)}) = \mathbb{E}\left\{\log \mathcal{L}(\boldsymbol{\beta}) | \boldsymbol{y}_{1:N}, \hat{\boldsymbol{\beta}}^{(k)}\right\}$
2. Compute $\hat{\boldsymbol{\beta}}^{(k+1)} = \arg\max_{\boldsymbol{\beta}} \mathcal{Q}(\boldsymbol{\beta},\hat{\boldsymbol{\beta}}^{(k)})$
3. Set $k \to k+1$ and return to Step 1.

For our problem, the $\mathcal{Q}$ function is given by

$$\mathcal{Q}(\boldsymbol{\beta},\hat{\boldsymbol{\beta}}^{(k)}) = \underbrace{\mathcal{Q}_1(\boldsymbol{\beta}_1,\hat{\boldsymbol{\beta}}^{(k)})}_{\text{Only depends on } \boldsymbol{x}_1} + \underbrace{\sum_{t=1}^{N} \mathbb{E}\left\{\log p(\boldsymbol{x}_{t+1}, y_t | \boldsymbol{x}_t, \boldsymbol{\beta}_2) | \boldsymbol{y}_{1:N}, \hat{\boldsymbol{\beta}}^{(k)}\right\}}_{\text{Depends on } \boldsymbol{\beta}_2 = (\boldsymbol{\Theta},\boldsymbol{\mu},\boldsymbol{\Sigma},\boldsymbol{a},\boldsymbol{b})},$$

where $\boldsymbol{x}_{t+1}, y_t | \boldsymbol{x}_t, \boldsymbol{\beta}_2 \sim \mathcal{TN}(\boldsymbol{\Theta}\boldsymbol{\zeta}_t + \boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Theta}\boldsymbol{\zeta}_t + \boldsymbol{a}, \boldsymbol{\Theta}\boldsymbol{\zeta}_t + \boldsymbol{b})$. Due to space constraints, we assume $\boldsymbol{x}_1$ known.

Theorem: Define $\mathcal{A}_t := \text{supp}\big(p(\boldsymbol{x}_t, \boldsymbol{x}_{t+1} | \boldsymbol{y}_{1:N}, \hat{\boldsymbol{\beta}}^{(k)})\big)$. The next EM iteration of $\boldsymbol{\beta}_2$ given $\hat{\boldsymbol{\beta}}^{(k)}$ is obtained by

$$\hat{\boldsymbol{\beta}}_2^{(k+1)} = \arg\max_{\boldsymbol{\beta}_2} \mathcal{V}_2^{(k)}(\boldsymbol{\beta}_2)$$

$$\text{s.t. } \boldsymbol{a} \leq \boldsymbol{\xi}_t - \boldsymbol{\Theta}\boldsymbol{\zeta}_t \leq \boldsymbol{b}, \quad (\boldsymbol{x}_t, \boldsymbol{x}_{t+1}) \in \mathcal{A}_t, t = 1, \ldots, N,$$

where

$$\mathcal{V}_2^{(k)}(\boldsymbol{\beta}_2) = \frac{-1}{2}\text{tr}\left\{\boldsymbol{\Sigma}^{-1}\left(\begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{\Theta}^\top \end{bmatrix}^\top \begin{bmatrix} \boldsymbol{\Xi} - \boldsymbol{\Phi}\boldsymbol{\Phi}^\top & -\boldsymbol{\Psi} + \boldsymbol{\Phi}\boldsymbol{\Lambda}^\top \\ -\boldsymbol{\Psi}^\top + \boldsymbol{\Lambda}\boldsymbol{\Phi}^\top & \boldsymbol{\Delta} - \boldsymbol{\Lambda}\boldsymbol{\Lambda}^\top \end{bmatrix} \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{\Theta}^\top \end{bmatrix} + (\boldsymbol{\Phi} - \boldsymbol{\Theta}\boldsymbol{\Lambda} - \boldsymbol{\mu})(\%)^\top\right)\right\} - \log \int_{\boldsymbol{a}}^{\boldsymbol{b}} \exp\left\{\frac{(\boldsymbol{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}{-2}\right\} \text{d}\boldsymbol{x},$$

and where we have defined the following quantities

$$\begin{bmatrix} \boldsymbol{\Xi} & \boldsymbol{\Psi} & \boldsymbol{\Phi} \\ \boldsymbol{\Psi}^\top & \boldsymbol{\Delta} & \boldsymbol{\Lambda} \\ \boldsymbol{\Phi}^\top & \boldsymbol{\Lambda}^\top & 1 \end{bmatrix} = \frac{1}{N}\sum_{t=1}^{N} \int \begin{bmatrix} \boldsymbol{\xi}_t \\ \boldsymbol{\zeta}_t \\ 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{\xi}_t \\ \boldsymbol{\zeta}_t \\ 1 \end{bmatrix}^\top p(\boldsymbol{x}_{t+1}, \boldsymbol{x}_t | \boldsymbol{y}_{1:N}, \hat{\boldsymbol{\beta}}^{(k)}) \text{d}\boldsymbol{x}_{t+1} \text{d}\boldsymbol{x}_t.$$

How to compute $\boldsymbol{\Xi},\boldsymbol{\Psi},\boldsymbol{\Phi},\boldsymbol{\Lambda},\boldsymbol{\Delta}$? Particle smoothing at each EM iteration (also required for computing $\mathcal{A}_t$).

How to maximize $\mathcal{V}_2^{(k)}(\boldsymbol{\beta})$? We require two lemmas:

Lemma 1: For fixed $(\hat{\boldsymbol{\Theta}}, \hat{\boldsymbol{a}}, \hat{\boldsymbol{b}})$, the $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ parameters that maximize $\mathcal{V}_2^{(k)}(\boldsymbol{\beta})$ satisfy

$$\hat{\mathcal{M}}_1 = \boldsymbol{\Phi} - \hat{\boldsymbol{\Theta}}\boldsymbol{\Lambda} - \boldsymbol{\mu},$$

$$\hat{\mathcal{M}}_2 - \hat{\mathcal{M}}_1\hat{\mathcal{M}}_1^\top = \begin{bmatrix} \boldsymbol{I} \\ \hat{\boldsymbol{\Theta}}^\top \end{bmatrix}^\top \begin{bmatrix} \boldsymbol{\Xi} - \boldsymbol{\Phi}\boldsymbol{\Phi}^\top & -\boldsymbol{\Psi} + \boldsymbol{\Phi}\boldsymbol{\Lambda}^\top \\ -\boldsymbol{\Psi}^\top + \boldsymbol{\Lambda}\boldsymbol{\Phi}^\top & \boldsymbol{\Delta} - \boldsymbol{\Lambda}\boldsymbol{\Lambda}^\top \end{bmatrix} \begin{bmatrix} \boldsymbol{I} \\ \hat{\boldsymbol{\Theta}}^\top \end{bmatrix},$$

where $\hat{\mathcal{M}}_i$ is the ith moment of the truncated Gaussian $\mathcal{TN}(\boldsymbol{0}, \boldsymbol{\Sigma}, \hat{\boldsymbol{a}} - \boldsymbol{\mu}, \hat{\boldsymbol{b}} - \boldsymbol{\mu})$.

Lemma 2: For fixed $(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}})$, the cost $\mathcal{V}_2^{(k)}(\boldsymbol{\beta})$ is maximized by

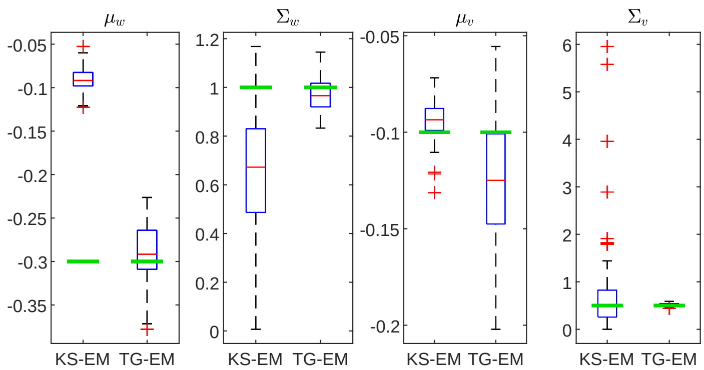
$$(\text{vec}\{\hat{\boldsymbol{\Theta}}\}, \hat{\boldsymbol{a}}, \hat{\boldsymbol{b}}) = \arg\max_{\boldsymbol{z},\boldsymbol{a},\boldsymbol{b}} \boldsymbol{z}^\top \text{vec}(\hat{\boldsymbol{\Sigma}}^{-1}[\boldsymbol{\Psi} - \hat{\boldsymbol{\mu}}\boldsymbol{\Lambda}^\top]) - \frac{1}{2}\boldsymbol{z}^\top(\boldsymbol{\Delta} \otimes \hat{\boldsymbol{\Sigma}}^{-1})\boldsymbol{z}$$

$$\text{s.t. } \begin{bmatrix} \boldsymbol{\zeta}_t \otimes \boldsymbol{I} \\ -\boldsymbol{\zeta}_t \otimes \boldsymbol{I} \end{bmatrix} \boldsymbol{z} \leq \begin{bmatrix} \boldsymbol{\xi}_t - \boldsymbol{a} \\ \boldsymbol{b} - \boldsymbol{\xi}_t \end{bmatrix}, \quad (\boldsymbol{x}_t, \boldsymbol{x}_{t+1}) \in \mathcal{A}_t, t = 1, \ldots, N.$$

Implementation aspects? fixed-point iterations + control variates + quadratic programming

Simulation study

Setup: $(\boldsymbol{A},\boldsymbol{B},\boldsymbol{C},D)$ fixed, $\boldsymbol{\mu} = [-0.3, -0.1]$, $\boldsymbol{\Sigma} = \text{diag}\{1, 0.5\}$, $\boldsymbol{a} = [-1.5, -50]^\top$, $\boldsymbol{b} = [2.5, 50]^\top$, $N = 5000$.



Reference

[1] R. A. González, A. L. Cedeño, K. Tiels and T. Oomen. "Identification of Linear State-Space Models Subject to Truncated Gaussian Disturbances". In preparation.